The Murine Testicular Transcriptome: Characterizing Gene Expression in the Testis During the Progression of Spermatogenesis

James E. Shima, Derek J. McLean, John R. McCarrey, and Michael D. Griswold

Center for Reproductive Biology, School of Molecular Biosciences, Washington State University, Pullman, Washington 99164
Department of Animal Sciences, Washington State University, Pullman, Washington 99164
Department of Biology, University of Texas, San Antonio, Texas 78249

ABSTRACT

One of the most promising applications of microarrays is the study of changes in gene expression associated with the growth and development of mammalian tissues. The testis provides an excellent model to determine the ability of microarrays to effectively characterize the changes in gene expression as an organ develops from birth to adulthood. To this end, a developmental testis gene expression time course profiling the expression patterns of ∼36,000 transcripts on the Affymetrix MGU74v2 GeneChip platform at 11 distinct time points was created to gain a greater understanding of the molecular changes necessary for and elicited by the development of the testis. Additionally, gene expression profiles of isolated testicular cell types were created that can aid in the further characterization of the specific functional actions of each cell type in the testis. Statistical analysis of the data revealed 11,252 transcripts (9846 unique) expressed differentially in a significant manner. Subsequent cluster analysis produced five distinct expression patterns within the time course. These patterns of expression are present at distinct chronological periods during testis development and often share similarities with cell-specific expression profiles. Analysis of cell-specific expression patterns produced unique and characteristic groups of transcripts that provide greater insight into the activities, biological and chronological, of testicular cell types during the progression of spermatogenesis. Further analysis of this time course can provide a distinct and more definitive view into the genes implicated, known and unknown, in the maturation, maintenance, and function of the testis and the integrated process of spermatogenesis.

INTRODUCTION

The testes are part of a select set of organs necessary for the continued propagation of a species that reproduces sexually. Specifically, testes are the site of spermatogenesis, a complex process in which the somatic Sertoli cells of the seminiferous epithelium support the differentiation of germ cells into functionally competent spermatozoa. The production of spermatozoa requires precise and highly ordered gene expression in order to initiate and maintain the numerous cellular processes leading to mature spermatozoa, including various mitotic, meiotic, and differentiation events [1]. Gene expression in the testis can be regulated by a variety of factors, including hormones, cell-cell interactions, and environmental cues, each of which can occur simultaneously or at distinct periods during the maturation of the animal leading to sexual maturity. Identification of the genes involved in the production of sperm via spermatogenesis has shown that both unique and ubiquitous genes play key roles in this process [1]. However, identifying and characterizing the genes necessary for spermatogenesis and their expression patterns at the genomic level has proven difficult due to the complex nature of an organ consisting of numerous cell types, each contributing to the total testicular transcriptome.

Once the first wave of spermatogenesis has been completed, spermatozoa are produced asynchronously throughout postpubertal life, making identification of genes involved in specific actions or from specific cell types within the testis difficult to pinpoint [2]. However, during the prepubertal period of development and continuing through the first wave of spermatogenesis, germ cells multiply and differentiate in a synchronous fashion in the testes [3]. Therefore, the analysis of changes in gene expression based on major stages of development, especially in regard to germ cell development, will provide a powerful tool to determine the cellular processes involved in the formation of spermatozoa.

Microarrays provide a platform to evaluate the abundance of thousands of genes using a single biological sample [4–6]. The application of microarrays to investigate gene expression in a single cell suspension in vitro over a time course following hormone treatment demonstrated the applicability of this technique to the analysis of hormonal regulation of gene expression [7]. Microarray studies have also been undertaken to determine the expression of testis or germline related genes in invertebrate animals such as Drosophila melanogaster [8] and Caenorhabditis elegans [9]. The analysis of gene expression using microarrays during organ development in vertebrates has additional challenges due to the changes in cell populations, the timing of cell differentiation, the variation of gene expression due to changing cell activity, and the general increase in complexity of cellular activities in higher eukaryotes. This is particularly true in the mammalian testis, in which the somatic cells and the differentiating germ cells undergo dramatic changes in transcription and differentiation. The complexity of the differentiating testis provides an interesting model to determine if microarray analysis of gene expression is adequate to decipher the changes occurring in the transcriptome. Furthermore, the timing of the appearance, prolifer-
and differentiation of individual cell types in the testis, and many of the genes associated with these cell types has been firmly established in the mouse, thus providing a powerful means to determine if microarray data recapitulate prior knowledge of the system. Application of this knowledge has many implications in fertility treatments and new contraceptive options.

The mouse testis developmental testis time course was designed to resolve the expression profile of approximately 36,000 transcripts throughout the development, taking into consideration the simultaneous progression of spermatogenesis. The ultimate goal of this time course was characterizing the expression profiles of the testis as a whole and that of the individual cell types that comprise it. Analysis of the time course confirmed that microarray data recapitulates known gene expression patterns in the testes at the organ and cellular level. Furthermore, the genes involved in numerous steps of spermatogenesis, their relative abundance, and the number of genes with changing expression patterns during major stages of development were also determined. Cell-specific gene expression was also established using enriched testicular cell populations of Sertoli, myoid, and germ cells, including type A and B spermatogonia, pachytene spermatocytes, and round spermatids. These expression profiles further emphasize the known functional activities of each cell type and present new possibilities for other functional processes and the involvement of previously unknown genes in the maturation of the testis and differentiation of germ cells.

Although not conclusive by nature, characterization of the testis transcriptome in this fashion effectively describes the major changes during testis development and the major transcripts present in each cell type in the testis and provides a useful resource for future studies of the biological factors influencing testis development and spermatogenesis.

**MATERIALS AND METHODS**

**RNA Preparation**

Protocols for the use of animals in this experiment were approved by the Washington State University and University of Texas at San Antonio Animal Care and Use Committees and were in accordance with the National Institute of Health’s standards established by the Guidelines for the Care and Use of Experimental Animals. A BL6-129 mouse colony was maintained in a temperature- and humidity-controlled room with food and water provided ad libitum. Whole testes were obtained from neonates at Days 0, 3, 6, 8, 10, 14, 18, 20, 30, 35, and 56 postpartum. Replicate samples were obtained in a similar fashion from the same breeding pairs.

Mouse Sertoli and peritubular myoid cells were obtained from 16- to 18-day-old whole testes from BL6-129 mice using a primary culture system and were cultured for a minimum of 2 days at 37°C with 5% CO₂ using standard culture procedures in Dulbecco modified Eagle medium and Ham F12 media following the method described by Karl and Griswold [10]. Enriched germ cell isolates were obtained from CD1 mice from the McCarey lab at the University of Texas, San Antonio, via gravity sedimentation [11]. These samples included type A and B spermatogonia, pachytene spermatocytes, and round spermatids from animals 8, 30, and 60 days of age, respectively. The purity of all the isolated and cultured cells was determined via morphological criteria. The purity of type A and B spermatogonia was >85% and the purity of all other testicular cell types was >95%.

Additionally, adult mouse brain and liver samples and whole testis samples from W/W v mice were obtained and used in the analysis. Brain and liver samples were used in the subtraction process in order to eliminate high-abundance transcripts that tended to be ubiquitous throughout mouse tissues. The W/W' testis sample, which, due to a mutation in the kit receptor (c-kit) has a negligible number of germ cells, provided a closer in vivo model to the whole testis and was used to isolate specific germ cell expression profiles rather than subtracting each individual somatic cell type. The W/W' sample was also used to indirectly determine the expression profile of Leydig cells.

All of the excised testicular and tissue samples were immediately placed in TRIzol reagent (Invitrogen, Carlsbad, CA) and stored at −70°C until total RNA was extracted according to the manufacturer’s standard protocol. cDNA from the isolated testicular cell types was extracted using the same protocols from the manufacturer. RNA quality was determined by both electrophoretic methods using a denaturing agarose gel and absorption readings at 260 and 280 nm. If excessive degradation or protein contamination of the RNA was evident, the sample was not used for microarray hybridization. A minimum optical density (OD260/280) ratio of 1.8 was required for microarray hybridization.

**Microarray Processing**

Ten micrograms of total RNA from each of the testicular and tissue samples was used to create the target for the microarray. Using the terminology of this microarray system, target refers to biotinylated complementary RNA (cRNA) created from the total RNA while probe refers to the oligonucleotides synthesized on the microarray platform. The biotinylated cRNA was generated using an oligo(dT) primer with a T7 promoter in a reverse transcription reaction, followed by in vitro transcription using the MEGAScript kit (Ambion, Austin, TX) with biotinylated cytosine and uridine triphosphate. The labeled cRNA was fragmented, hybridized to MGU74Av2, Bv2, and Cv2 arrays (Affymetrix, Santa Clara, CA), and stained in accordance with the manufacturer’s standard protocol. The same procedure was performed on duplicate samples with the exception of liver, brain, W/W’ testis, and germ cell samples, which had only one sample. The arrays were washed and stained utilizing the Affymetrix GeneChip Fluidics Station 400 and scanned using a GeneArray Scanner 2500A (Agilent, Palo Alto, CA). The resulting data were viewed and preliminary assessment was made using Microarray Suite 5.0 (MAS) software (Affymetrix). All reactions and microarray hybridization procedures were performed in the Laboratory for Biotechnology and Bioanalysis (LBBI) at Washington State University.

**Absolute and Statistical Analysis**

Microarray output was examined visually for excessive background noise and physical anomalies. The default MAS statistical values were used for all analyses. All probe sets on each array were scaled to a mean target signal intensity of 125, with the signal correlating to the amount of transcript in the sample. An absolute analysis using MAS was performed to assess the relative abundance of the −36,000 represented transcripts based on signal and detection (present, absent, or marginal). Although −36,000 probe sets are represented on these arrays, there is a certain level of redundancy within and between each array, which results in a lower number of unique transcripts represented on the MGU74v2 chipset; however, the level of redundancy has not been firmly established.

The absolute analysis from MAS was imported into GeneSpring 5.1 software (Silicon Genetics, Redwood City, CA). The developmental time course data was normalized within GeneSpring using the default/recommended normalization methods. These include setting of signal values below 0.01 to 0.01, total chip normalization to the 50th percentile, and normalization of each gene to the median. Isolated cell types, liver, brain, and W/W’ testis samples, were normalized in the same manner but without normalizing each gene to the median. These normalizations allowed for the visualization of data based on relative abundance at any given time point rather than compared to a specific control value.

Data restrictions and analytical tools in GeneSpring were applied to isolate noteworthy and possibly important patterns of gene expression during the course of testicular development and spermatogenesis. Transcripts expressed differentially at a statistically significant level were determined using a one-way ANOVA parametric test with variances not assumed equal, a P-value cutoff of 0.05 and using a Benjamini and Hochberg False discovery rate multiple testing correction. This was applied to all 11 time points and considered all transcripts represented on the arrays. Subsequently expression restrictions were applied to the transcripts expressed in a significant manner. These restrictions were designed so that the remaining transcripts met the following requirements in addition to being expressed in a significant manner: 1) each transcript must have a signal value of at least 100 in a minimum of 2 out of 11 points in the time course and 2) the range of the replicates must not exceed 1 (in the normalized scale) in half of the time points comprising the time course. The resulting transcripts were screened using Excel (Microsoft, Redmond, WA) for redundant UniGene entries. Transcripts that passed these restrictions were considered for further analysis.
Gene Expression in the Developing Murine Testis

Cluster Analysis

Patterns of gene expression were identified using unsupervised cluster analysis within the set of differentially expressed transcripts that met the requirements detailed previously. Clustering algorithms allow for the separation of distinct patterns of expression based on the similarity of expression profiles between different genes [12]. In this analysis, a hierarchical clustering algorithm utilizing a smooth correlation with the default parameters was utilized in order to isolate distinct, nonrepetitive patterns of expression within the time course. A nonphylogenetic gene tree was produced from this analysis that illustrates the major expression patterns within the differentially expressed transcripts (determined through statistical analysis) in a continuous fashion.

In Silico Subtraction

Using a method based on in silico subtraction techniques, two categories describing the localization of transcripts within the testis can be made: 1) transcripts can either be expressed at a high level compared with other cell types in the testis (enriched) or 2) transcripts can be expressed uniquely in a particular cell type compared with other cell types in the testis (specific). These characteristics were made within GeneSpring based on the raw data associated with each transcript from each tissue and cell experiment of interest. Enriched transcripts were required to have a signal level of at least 200 and be at least 2-fold higher in the cell type of interest than in any other cell or tissue sample. Specific transcripts had to meet the enriched specifications and have signal levels below 100 in all other cell and tissue types. As an example, each transcript determined to be enriched in Sertoli cells had to have a signal level of at least 200 in the Sertoli cell microarray experiment and that signal level had to also be 2-fold higher than in every other experimental dataset including all four germ cell types, myoid cells, and brain and liver tissue. Transcripts considered Sertoli cell specific were required to have a signal level below 100 in all other cell and tissue types as well as meet the enriched transcript requirements. Expression profiles were identified for Sertoli, Leydig, myoid, type A and B spermatagonia, pachyteme spermatocytes, and round spermatids using these parameters. W/W−/− testis samples were used in the subtractions for all germ cell types and for indirect Leydig cell-enriched spermatids using these parameters. W/W−/− testis samples were used in the subtractions for all germ cell types and for indirect Leydig cell-enriched spermatids using these parameters. W/W−/− testis samples were used in the subtractions for all germ cell types and for indirect Leydig cell-enriched spermatids using these parameters.

RESULTS

Global Analysis

A genome-wide transcriptional analysis of the developing murine testis and spermatogenesis was done using the Affymetrix GeneChip system. Approximately 36,000 transcripts, representing the entire mouse genome (based on build 74 of the UniGene database) were monitored from birth (Day 0) through adulthood (Day 54). In all, 66 arrays (two independent replicates per time point using the MGU74v2 chipset. When these transcripts were displayed in a parallel-coordinate plot, a pattern of expression was evident wherein a high level of expression of numerous transcripts was evident early in testis development that tapered off to a significantly lower level prior to meiosis and the appearance of pachytenic spermatocytes approximately 14 days after birth. This was followed by elevated expression of an altogether unique set of transcripts through to adulthood, as shown in Figure 1. This pattern illustrates the importance of the major periods in testicular development: the initial formation of an environment conducive to spermatogenesis, the entrance into meiosis, and resulting production and proliferation of haploid gametes within the testis.

Cluster Analysis

Cluster analysis was used to better characterize the more specific and unique patterns of expression of the 9846 unique and statistically significant transcripts. A hierarchical clustering algorithm was utilized within GeneSpring to generate a gene tree representative of the major patterns of expression occurring in this group of transcripts (Fig. 2). Five distinct patterns with relatively discrete chronological boundaries emerged from within the time course. The five unique patterns included high expression from birth to 6–8 days postpartum (Fig. 2A), an elongated period of elevated expression between Day 0 and 14 days postpartum (Fig. 2B), induction at Days 6–8, with repression occurring

### Table 1. Transcripts expressed at various stages during the first wave of spermatogenesis.

<table>
<thead>
<tr>
<th>Day postpartum</th>
<th>Percentage present</th>
<th>Number present</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>34.3</td>
<td>12,723</td>
</tr>
<tr>
<td>3</td>
<td>33.4</td>
<td>12,397</td>
</tr>
<tr>
<td>6</td>
<td>33.6</td>
<td>12,447</td>
</tr>
<tr>
<td>8</td>
<td>33.5</td>
<td>12,408</td>
</tr>
<tr>
<td>10</td>
<td>36.1</td>
<td>13,395</td>
</tr>
<tr>
<td>14</td>
<td>37.3</td>
<td>13,812</td>
</tr>
<tr>
<td>18</td>
<td>34.5</td>
<td>12,787</td>
</tr>
<tr>
<td>20</td>
<td>31.8</td>
<td>11,772</td>
</tr>
<tr>
<td>30</td>
<td>32.5</td>
<td>12,043</td>
</tr>
<tr>
<td>35</td>
<td>30.4</td>
<td>11,260</td>
</tr>
<tr>
<td>56</td>
<td>29.2</td>
<td>10,822</td>
</tr>
</tbody>
</table>
between Days 14 and 18 (Fig. 2C), dramatic upregulation beginning between 14 and 18 days after birth and remaining elevated until adulthood (Fig. 2D), and dramatic upregulation between Days 20 and 30 with sustained elevated levels through to adulthood (Fig. 2E). These specific groups contained 1169, 3322, 1027, 3210, and 1118 transcripts, respectively, with no overlap between the transcripts comprising each group. Although not definitive, the presence of previously identified genes within each of these groups may infer similar roles for the other transcripts present in the same clusters, such as meiosis in cluster C and D and the specialized structure and function of spermatozoa in clusters D and E. A list of genes essential to spermatogenesis, as determined through knockout models, is presented in Table 2. These transcripts are indicative of some of the processes that are occurring at various stages of spermatogenesis and further demonstrate how genes required for complete spermatogenesis exhibit varied expression patterns and levels throughout testis development. Complete lists of transcripts from each cluster will be available as supplemental data through the Griswold Lab homepage (http://www.wsu.edu/~griswold/microarray).

Testicular Cell-Type Expression Profiling

Gene expression profiles of testicular cell types were obtained through in silico subtraction using microarray results from isolated cell type samples. The groups of transcripts defined using this technique can serve as a starting point for further investigation into cell-specific transcripts and their actions within each specific cell type and the testis. The number of transcripts considered enriched and specific in each cell type varied significantly, with round spermatids possessing the greatest number of enriched and specific transcripts while spermatogonia exhibited the smallest number of enriched and specific transcripts. This reinforces previous findings demonstrating that advanced germ cells (specifically spermatocytes and spermatids) possess a large number of genes with unique functions and specific localization to the testis. Somatic cells present in the testes presented a varied range of enriched and specific transcripts, the numbers of which fall between the levels of spermatogonia and more advanced germ cells. The number of enriched and specific transcripts from each cell type is shown in Table 3. The results of these subtractions yielded groups of transcripts that exhibited unique patterns of expression when plotted in the testis developmental time course (using the data associated with mRNA levels present in a complete testis rather than isolated cell types). Analysis within the context of the developmental time course also allows for the identification of unique expression patterns of individual transcripts during testis development based on testicular cell type. Complete expression profiles of enriched transcripts in each cell type are shown in Figures 3 and 4. These figures represent the general pattern of expression of the major testicular cell types during testis development. As expected, advanced meiotic and postmeiotic germ cells exhibited a high degree of induction of transcripts beginning at 14 days postpartum and 20 days postpartum, respectively, with the vast majority of these transcripts being maintained at an elevated level through to adulthood, as indicated in Figure 3. A limited number of significant changes in expression of these transcripts were detected prior to the appearance of spermatocytes at 10–14 days postpartum and spermatids at about 18 days postpartum. Transcripts enriched in the two types of spermatogonia demonstrated an opposing pattern of expression, with distinct peaks in number and expression level occurring both perinatally and 10–14 days postpartum, although some significant levels of expression did occur during the later stages of spermatogenesis between 14 and 56 days postpartum. Sertoli-, Leydig-, and myoid cell-enriched transcripts exhibited high levels of expression between birth and 6 days postpartum, while Sertoli cell-enriched transcripts also had a distinctive pattern of gene regulation similar to that found in cluster C of Figure 2, in which transcripts were induced approximately 8 days postpartum, reached a peak level of expression by 10–14 days postpartum, and were returned to basal levels by 14–18 days postpartum. Each of the so-
GENE EXPRESSION IN THE DEVELOPING MURINE TESTIS

FIG. 2. Clustering of transcripts expressed at a statistically significant level during testis development. The gene tree shows the predominant patterns of gene expression of 9846 significantly expressed transcripts during spermatogenesis. Clustering was done within GeneSpring 5.1 using a smooth correlation with the default separation ratio and minimum distance. Letters A-E indicate distinct patterns of expression identified through clustering.

Each of these expression profiles recapitulated previously described and expected biological processes occurring in each cell type and at specific time frames in testicular development. Some of these processes are evident, as shown in Table 4, which relates previously known genes with specificity to certain cell types. In particular, Leydig cell-specific transcripts exhibited a number of genes involved in steroidogenesis and DNA condensation and sperm-structure-related genes were found in the meiotic and postmeiotic germ cell (pachytene spermatocytes and round spermatids)-specific transcript groupings. Complete lists of transcripts of enriched and specific transcripts will be available as supplemental data through the Griswold Lab homepage (http://www.wsu.edu/~griswold/microarray).

DISCUSSION

During the development of the testis and the associated process of spermatogenesis, well-ordered and sequential changes in gene expression are required in order to create a fully functional testis capable of producing mature spermatozoa. Previous work using knockout or transgenic models, analyzing the functional impact of a single gene, have shown that, when the time frame or sequence of gene expression is altered, fertility issues can arise [65, 66]. This is a time-consuming and labor-intensive approach and ultimately impractical for analysis of a group of genes as large as those in a complete mammalian genome. An alternate approach is to look at changes in gene expression at a global level in a wild-type organism, studying an entire organ and its different cell types undergoing a normal process of differentiation and growth rather than the sum of its parts. This has been done on numerous occasions by investigators examining gene expression changes under certain conditions or in specific cell types in vitro [7, 67, 68]. These studies, although informative and insightful, are limited in their abilities to describe the transcriptional changes occurring in an entire organ under in vivo conditions.

Recently, studies have been undertaken to better understand changes in gene expression during spermatogenesis using various knockout models in combination with genome-scale analysis. Maratou et al. utilized cDNA microarrays with wild-type and \textit{dazl} knockout mouse testes to examine the effects that \textit{dazl} has on spermatogenesis [69]. However, this study focused on ~1300 unique genes obtained from a testis cDNA library, limiting the scope of their analysis. Similarly, O’Shaughnessy and colleagues utilized serial analysis of gene expression (SAGE) to characterize gene expression in somatic cells of the testis by using W\textsuperscript{v}v/W\textsuperscript{v}v mutant testis samples [70]. These studies did not examine the majority of the genes expressed in the testis but do provide useful datasets to complement genome-wide, wild-type testis datasets, such as the one in the current study.

The approach utilized in this study provides the opportunity to examine thousands of candidate genes based on their expression profile in the context of a time course of wild-type testis development. The appearance of distinct germ cell types was used as reference points for the construction of the time course, thus allowing spermatogenesis,
TABLE 2. Transcripts required for spermatogenesis as determined through knockout models.*

<table>
<thead>
<tr>
<th>Description</th>
<th>GenBank no.</th>
<th>Cluster</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inhibin alpha</td>
<td>X55957</td>
<td>A</td>
<td>[13]</td>
</tr>
<tr>
<td>GDP dissociation inhibitor 1</td>
<td>U07950</td>
<td>A</td>
<td>[14]</td>
</tr>
<tr>
<td>Insulin-like growth factor 1</td>
<td>X04480</td>
<td>A</td>
<td>[15]</td>
</tr>
<tr>
<td>Jun proto-oncogene related gene d1</td>
<td>J04509</td>
<td>A</td>
<td>[16]</td>
</tr>
<tr>
<td>Kit ligand</td>
<td>M57647</td>
<td>B</td>
<td>[17]</td>
</tr>
<tr>
<td>Bcl2-associated X</td>
<td>L22472</td>
<td>B</td>
<td>[18]</td>
</tr>
<tr>
<td>Mitochondrial insertion mutation</td>
<td>Ali83160</td>
<td>B</td>
<td>[19]</td>
</tr>
<tr>
<td>Desert hedgehog</td>
<td>Ali666359</td>
<td>B</td>
<td>[20]</td>
</tr>
<tr>
<td>Solute carrier family 12, member 2</td>
<td>U13174</td>
<td>B</td>
<td>[21]</td>
</tr>
<tr>
<td>Claudin 11</td>
<td>U19582</td>
<td>C</td>
<td>[22]</td>
</tr>
<tr>
<td>Kit receptor</td>
<td>Y00864</td>
<td>C</td>
<td>[23]</td>
</tr>
<tr>
<td>Ataxia telangiectasia mutated homolog</td>
<td>U43678</td>
<td>C</td>
<td>[24]</td>
</tr>
<tr>
<td>Dax1</td>
<td>U41568</td>
<td>C</td>
<td>[25]</td>
</tr>
<tr>
<td>Double-sex and mab-3 related transcription factor 1</td>
<td>AL133300</td>
<td>C</td>
<td>[26]</td>
</tr>
<tr>
<td>Cyclin A1</td>
<td>X84311</td>
<td>D</td>
<td>[27]</td>
</tr>
<tr>
<td>Heat shock protein 70-2</td>
<td>L27086</td>
<td>D</td>
<td>[28]</td>
</tr>
<tr>
<td>TATA box binding protein-like 1</td>
<td>AB017697</td>
<td>D</td>
<td>[29]</td>
</tr>
<tr>
<td>Calmodulin</td>
<td>U08373</td>
<td>D</td>
<td>[30]</td>
</tr>
<tr>
<td>A disintegrin and metalloprotease domain 2</td>
<td>U16242</td>
<td>D</td>
<td>[31]</td>
</tr>
<tr>
<td>MutS homolog 4</td>
<td>AV277719</td>
<td>D</td>
<td>[32]</td>
</tr>
<tr>
<td>Transition protein 1</td>
<td>X12521</td>
<td>E</td>
<td>[33]</td>
</tr>
<tr>
<td>Transition protein 2</td>
<td>M60254</td>
<td>E</td>
<td>[34]</td>
</tr>
<tr>
<td>Protamine 1</td>
<td>X07625</td>
<td>E</td>
<td>[35]</td>
</tr>
<tr>
<td>Seven in absentia 1a</td>
<td>AA492653</td>
<td>E</td>
<td>[36]</td>
</tr>
<tr>
<td>Hormone-sensitive lipase</td>
<td>U69543</td>
<td>E</td>
<td>[37]</td>
</tr>
<tr>
<td>Phosphoserine/threonine/tyrosine interaction protein</td>
<td>U34973</td>
<td>E</td>
<td>[38]</td>
</tr>
<tr>
<td>Testicular haploid expressed gene</td>
<td>AJ011834</td>
<td>E</td>
<td>[39]</td>
</tr>
</tbody>
</table>

* Requirement for spermatogenesis based on knockout models listed by Matzuk and Lamb [40].

TABLE 3. Enriched and specific transcripts in each testicular cell type as determined through in silico subtraction.

<table>
<thead>
<tr>
<th>Testicular cell type</th>
<th>Enriched</th>
<th>Specific</th>
<th>Percentage with significant change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leydig</td>
<td>126</td>
<td>26</td>
<td>45</td>
</tr>
<tr>
<td>Peritubular myoid</td>
<td>437</td>
<td>85</td>
<td>53</td>
</tr>
<tr>
<td>Sertoli</td>
<td>152</td>
<td>30</td>
<td>59</td>
</tr>
<tr>
<td>Type A spermatogonia</td>
<td>50</td>
<td>4</td>
<td>62</td>
</tr>
<tr>
<td>Type B spermatogonia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pachytene</td>
<td>39</td>
<td>9</td>
<td>33</td>
</tr>
<tr>
<td>Spermatocytes</td>
<td>364</td>
<td>34</td>
<td>72</td>
</tr>
<tr>
<td>Round spermatids</td>
<td>694</td>
<td>215</td>
<td>87</td>
</tr>
</tbody>
</table>

the process in which germ cells multiply and differentiate into mature spermatozoa, to act as a guide to critically analyze the time course.

The time course allows for queries concerning the testicular expression of the majority of genes of the murine genome. However, limitations on the interpretation of the results of these queries should be acknowledged. It has been reported that the Affymetrix GeneChips may yield up to 2% false positives, although transcripts are detected consistently and accurately at or above a 4-pM concentration [71]. These studies also measure only steady-state mRNA levels and may not reflect protein levels. As such, induction or repression of a transcript may not relate directly to functional changes. Additionally, in the developmental time course, the dynamic changes in expression levels for a particular transcript may be affected dramatically by changing cell populations. For example, transcripts in Sertoli cells appear to decrease at a slow but steady rate after the initiation of meiosis, but this is likely due in part to the relatively large increase in germ cell messages during the latter period in the initial wave of spermatogenesis. As a result, the induction or repression of transcripts expressed in specific testicular cell types may not entirely be the result of direct upregulation or downregulation and may in fact be the indirect result of changes in the cell composition of the testis. Figures 3 and 4, which illustrate the cell-specific expression profiles of testicular cell types during testis development, therefore may also be indicative of regulation due in part to changes in cell populations, particularly during the meiotic and postmeiotic periods of development.

Cultured Sertoli and myoid cells and isolated germ cell samples were used to identify enriched and specific transcripts in this study. The removal of Sertoli cells from their normal in vivo environment has been shown to alter the expression of various genes by disrupting the normal signaling that occurs between Sertoli cells and germ cells [72–74] and between different germ cell populations [75], and it can be inferred that the isolation of myoid cells using a culture system would have similar effects. Plotting these transcripts using the data associated with the developmental time course provides the ability to identify transcripts altered by the in vitro conditions and thus compensate for the isolation procedures performed on these cells. We have attempted to minimize this problem further by employing duplicate arrays and using relatively high signal values as a cutoff for analysis.

Evaluation of the time course in regard to reliability and accuracy of the data is essential in order to make full use of the dataset and support any findings that result from its analysis. Therefore, genes previously studied in the testis were examined in order to gauge the relative accuracy of the data. This was accomplished through literature searches for genes expressed prominently in the testis and in the individual testicular cell types. The time course appears to reproduce the expected expression profiles of known testis.
FIG. 3. Parallel-coordinate plot of germ cell-enriched transcripts generated using GeneSpring 5.1. Transcripts were isolated using in silico subtraction techniques based on each cell type and replotted using the developmental time course data. Colors are representative of expression levels at Day 0 postpartum. X-axis is days postpartum and the y-axis is normalized intensity.

FIG. 4. Parallel-coordinate plot of somatic cell-enriched transcripts generated using GeneSpring 5.1. Transcripts were isolated using in silico subtraction techniques based on each cell type and replotted using the developmental time course data. Colors are representative of expression levels at Day 0 postpartum. X-axis is days postpartum and the y-axis is normalized intensity.
genes as determined through examination of the results of cluster analysis. Genes known to be required for or highly
expressed during meiosis, such as cyclin A1 [27] and heat shock protein 70-2 [28], as well as genes known to be as-
sociated with the postmeiotic germ cells, including prot-
amine 1 [35] and transition protein 1 [33] and 2 [34], were
found to be highly expressed at appropriate periods in the
time course, and their induction coincides with the expected
biological actions occurring within the testis. Genes nor-
mally expressed in the somatic cells of the testis, including
kit ligand [76] and clusterin in Sertoli cells [46], 3β-
hydroxysteroid dehydrogenase in Leydig cells [41], and cal-
cyclin in myoid cells [51] also reaffirms the dataset and our
ability to accurately identify genes with functional signif-
icance in the testis. The vast majority of genes examined fit
the expected profile within the time course both in induc-
tion and repression of expression and in peak signal levels.

The methods used to isolate cell-enriched and -specific transcripts were evaluated to determine the efficacy in iso-
lating said transcripts. Genes expressed in specific cell
types of the testis are shown in Table 4 and demonstrate
the ability to identify genes associated with certain cell
types using the devised in silico subtraction method. The
results also suggest that previously uncharacterized trans-
scripts found in these lists may also be enriched or
even specific in the respective cell types in the testis. How-
ever, the methods used to isolate these transcripts were rel-
atively stringent and, as a result many transcripts thought
to be specific in a particular cell type and supported by
experimental evidence failed to pass the restrictions applied
to isolate these groups of transcripts. This indicates that,
while useful, the in silico subtraction process employed to
isolate these transcripts is not without flaws. A possible
alternative approach for isolating cell-specific transcripts
would be to use statistical analysis rather than fold-change
cutoffs to identify transcripts based on significant differ-
ences in expression level between cell populations rather
than an arbitrary fold-change difference.

The overall pattern of gene expression during the period
of testicular development examined in this study demon-
strated three key features: 1) ~50% of the murine genome
(based on the Affymetrix microarray platform) may be in-
volved in testis development and/or spermatogenesis at any
point between birth and adulthood, 2) ~30% of the murine
genome is differentially expressed in a significant manner
during testis development, and 3) major periods of expres-
sional change occur during the first few days after birth (0–
6 days postpartum), at the initiation of meiosis (~14 days
postpartum), and with the appearance of haploid gametes
(~20 days postpartum). Limitations of this microarray plat-
form are evident and include incomplete coverage of the
murine genome (only 6000 well-characterized genes are
represented on the array set) and incorrectly represented
genes due to inaccurate probe sets, and this may play a role
in the number of transcripts contributing in testis function.
The percentage of the murine genome shown to be involved
in the testis is nonetheless indicative of the highly special-
ized role it plays in sperm production. A recent publication
using a similar time course and the same microarray plat-
form found that 21 274 (or ~58%) of the murine genome
was expressed during testis development in the mouse [64].
This corroborates our finding that a large portion of the
murine genome is possibly involved in testis formation and
function and spermatogenesis. Considering the unique
functional nature of Sertoli, peritubular myoid, and Leydig
cells as well as the specialized requirements associated with

<table>
<thead>
<tr>
<th>Common name</th>
<th>GenBack no.</th>
<th>Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leydig</td>
<td></td>
<td>Steroid biosynthesis</td>
<td>[41]</td>
</tr>
<tr>
<td>3β-HSD (type I)</td>
<td>M58567</td>
<td>Steroid biosynthesis</td>
<td>[42]</td>
</tr>
<tr>
<td>17β-HSD (type III)*</td>
<td>U66827</td>
<td>Steroid biosynthesis</td>
<td>[43]</td>
</tr>
<tr>
<td>Cyp17</td>
<td>M64863</td>
<td>Steroid biosynthesis</td>
<td>[43]</td>
</tr>
<tr>
<td>Vcam1</td>
<td>M84487</td>
<td>Cell adhesion</td>
<td>[44]</td>
</tr>
<tr>
<td>Sertoli</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gata1</td>
<td>X15763</td>
<td>Transcription factor</td>
<td>[45]</td>
</tr>
<tr>
<td>Clusterin</td>
<td>D14077</td>
<td>Secretory protein</td>
<td>[46]</td>
</tr>
<tr>
<td>Sox9*</td>
<td>A060869</td>
<td>Transcription factor</td>
<td>[47]</td>
</tr>
<tr>
<td>Claudin 11</td>
<td>U19582</td>
<td>Tight junction structure</td>
<td>[48]</td>
</tr>
<tr>
<td>Cathepsin L</td>
<td>X06086</td>
<td>Protease</td>
<td>[49]</td>
</tr>
<tr>
<td>Myoid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actin-a2, smooth muscle</td>
<td>X13297</td>
<td>Cytoskeletal component</td>
<td>[50]</td>
</tr>
<tr>
<td>Calcyclin</td>
<td>X66449</td>
<td>Cell cycle regulator</td>
<td>[51]</td>
</tr>
<tr>
<td>Collagen 4-a1</td>
<td>M15832</td>
<td>ECM structure</td>
<td>[52]</td>
</tr>
<tr>
<td>Collagen 4-a2</td>
<td>X04647</td>
<td>ECM structure</td>
<td>[52]</td>
</tr>
<tr>
<td>Premeiotic germ cells</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dazl</td>
<td>U46694</td>
<td>Developmental regulator</td>
<td>[53]</td>
</tr>
<tr>
<td>c-Jun*</td>
<td>X12761</td>
<td>Transcription factor</td>
<td>[54]</td>
</tr>
<tr>
<td>Tex20*</td>
<td>AA16823 2</td>
<td>Unknown</td>
<td>[55]</td>
</tr>
<tr>
<td>Meiotic and Postmeiotic germ cells</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pkg-2</td>
<td>M17299</td>
<td>Glycolytic enzyme</td>
<td>[56]</td>
</tr>
<tr>
<td>Ldhc*</td>
<td>M17587</td>
<td>Glycolytic enzyme</td>
<td>[57]</td>
</tr>
<tr>
<td>Pmr1*</td>
<td>X07625</td>
<td>Sperm DNA structure</td>
<td>[58]</td>
</tr>
<tr>
<td>Tenr</td>
<td>X84693</td>
<td>RNA binding</td>
<td>[59]</td>
</tr>
<tr>
<td>Zp3p38*</td>
<td>D17569</td>
<td>Zona pellucida binding</td>
<td>[60]</td>
</tr>
<tr>
<td>Gapds*</td>
<td>U09964</td>
<td>Glycolytic enzyme</td>
<td>[61]</td>
</tr>
<tr>
<td>Tp2*</td>
<td>M60254</td>
<td>Sperm DNA structure</td>
<td>[62]</td>
</tr>
<tr>
<td>CatSper1</td>
<td>A037647</td>
<td>Cation channel</td>
<td>[63]</td>
</tr>
</tbody>
</table>

* Indicates enriched genes also found to be specific.
meiosis, spermatogenesis, and the eventual fertilization re-
quirements of sperm, the amount of the genome dedicated
to and regulated in the testis seems reasonable and prob-
able.

The ultimate goal of any study utilizing microarrays to
examine gene expression on a genome-wide scale in a com-
plex tissue is to link the patterns of gene expression to the
functional significance and cellular composition of the tis-
ue of interest. Examination of wild-type gene expression
in the testis provides great insight into the overall expres-
sional changes necessary to elicit spermatogenesis and the
production of spermatozoa; however, determining the spe-
cific actions and the functional significance of these chang-
es in expression and relating that to the specific testicular
cell types is far more difficult. The goal of this study is an
attempt at profiling the transcriptome of each cell type and
relating that to the overall patterns of expression occurring
in the testis. Examination of the timing of gene expression
in each cell type based on overall expression patterns pre-
sented unique insights into the functional activities of each
cell type during specific periods of testis development.

Expression of enriched transcripts in Sertoli cells is pri-
marily focused at 10 days postpartum, with very few tran-
scripts showing elevated expression in the postmeiotic stag-
es of spermatogenesis and only slightly more expressed at
an elevated level perinatally. Many of these transcripts are
expressed similarly to the transcripts found in Figure 2, B
and C. Nearly 45% of the enriched transcripts for Sertoli
cells match these expression profiles, indicating that a large
portion of Sertoli cell-specific transcripts are differentially
expressed around the time of meiosis within the initiation
of meiosis at 10 days postpartum and the appearance of
pachytene spermatocytes at 14 days postpartum. Little to
no Sertoli cell-specific gene expression was observed early
or late in testis development. The similarity of the tran-
scripts expressed in an exclusive fashion in Sertoli cells to
those in the two patterns elucidated through cluster analysis
suggests that many of the Sertoli cell-enriched and -specific
transcripts are involved in the adaptation of these cells to
accommodate the newly forming meiotic and postmeiotic
germ cells. The presence of claudin 11, which has previ-
ously been found to be Sertoli cell specific, further empha-
sizes this possibility, considering its necessity for the com-
pletion of spermatogenesis and resultant fertility [22].

Type A and B spermatogonia demonstrated a similar ex-
pression profile to that of Sertoli cells, although slight dif-
fences were evident. Both types of spermatogonia had a
large percentage of enriched transcripts overlap with the
transcripts identified in clusters A and B (46%), indicating
a greater period of expression activity earlier in develop-
ment. Surprisingly, very little spermatogonial-specific ex-
pression was detected after meiosis had begun. The lack of
differential expression in spermatogonia during and after
meiosis does not necessarily indicate that there are no
changes occurring, but rather that these changes are based
on the expression of genes that are not unique to sper-
matogonia. The high level of overlap between spermatog-
onal enriched genes and the pattern of expression in clus-
ter B (elevated expression from birth to ~10 days postpar-
tum) suggests that many unique spermatogonial genes are
downregulated as germ cells enter meiosis. A specific tran-
script demonstrating a unique expression profile indicative
of localization to spermatogonia was dlk1 (delta-like ho-
molog 1), a gene for which the activity and localization in
the testis was previously unreported but has been implicat-
ed in inhibition of adipocyte differentiation [77] and car-
diomyocyte differentiation [78] and is highly expressed in
neuroblastoma cell lines [79]. Dlk1 was highly expressed
in Type A and B spermatogonia with nearly total loss of ex-
pression by 20 days postpartum and the appearance of
haploid germ cells. Its localization to type B spermatogonia
through in silico subtraction indicates its possible involve-
ment in the differentiation of spermatogonia and its rapid
decline in expression beginning at the onset of meiosis
further the possibility that it may be involved in the dif-
ferentiation that occurs that allows spermatogonia to pro-
gress in spermatogenesis and eventually enter meiosis. A
variant form of dlk1, fetal antigen-1 (FA1), was previously
localized to Leydig cells but lacks the membrane-bound C-
terminal region as well as the signal peptide, indicating that
the dlk1 transcript possibly has multiple functions in the
testis in alternatively spliced forms [80].

Expression early in testis development was predominat-
ed by myoid cell-enriched transcripts. Nearly 50% of all
myoid-enriched transcripts demonstrated a similar expres-
sion profile to that in clusters A and B, signifying a large
role in testis development associated with the early post-
natal stages of growth. Interestingly, myoid cells exhibited
the greatest number of enriched and specific transcripts of
the somatic testicular cell types examined by a wide mar-
gin. The high number of transcripts expressed exclusively
in myoid cells indicates that their role in testis function and
the seminiferous tubule environment may be greater than
originally thought. Many of these transcripts are associated
with the extracellular matrix and basal lamina of a cell,
including numerous collagenas and proteoglycans. These in-
clude collagen 4 alpha 1 and 2, which have previously been
shown to be specifically expressed in myoid cells [52]. The
expression of numerous genes related to the extracellular
matrix and basal lamina is indicative of the large role
that myoid cells play in establishing the gross morphological
structure of the seminiferous tubules. Recently, however,
publications have shown that proper myoid cell structure
and function is required for both proper testis formation
[81] and for the progression of spermatogenesis [82].
Therefore, many of the transcripts localized to myoid cells
may in fact be necessary for more than just the establish-
ment of connective tissue structures between the cells of
the testis.

Although no purified populations of Leydig cells were
available for this study, Leydig cell gene expression was
simulated using the W/W^v testis expression profile as a
starting point and subtracting the expression profiles of my-
oid and Sertoli cells and spermatogonia. Thus, the results
should be predominated by transcripts expressed in Leydig
cells. This was confirmed to some degree by the presence
of genes encoding two of the required sex steroid enzymes
necessary for the conversion of cholesterol to androgens
including, 3beta-hydroxysteroid dehydrogenase (3beta-HSD)
and 17beta-hydroxysteroid dehydrogenase (17beta-HSD). In addi-
tion, vascular cell adhesion molecule 1 (VCAM1), a known
Leydig cell-specific gene [44], was also found in the en-
riched group of transcripts obtained through in silico sub-
traction. This would tend to support our technique to sim-
ulate Leydig cell gene expression using the W/W^v testis
and in silico subtractions. The overall expression profile of
Leydig cells did not fit any profile achieved through cluster-
ing nor did it follow any similar pattern of other testic-
ular cell types. However, a small number of transcripts did
demonstrate a similar expression pattern to transcripts
found in cluster A. These transcripts could be involved in
early Leydig cell differentiation, from fetal to adult, which
occurs by 7 days postpartum in the mouse [83]. Many of these transcripts could be involved in this differentiation process and the conversion of luteinizing hormone-dependent cells to dependent cells.

The most significant changes in expression throughout the time course were observed after the initiation of meiosis and the subsequent appearance of spermatocytes and spermatids. As expected, the pachytene spermatocyte and round spermatid expression profiles were also >70% similar to clusters D and E, respectively. This is demonstrated by the expression profiles of pachytene spermatocytes and round spermatids in Figure 3, in which numerous transcripts exhibit a high degree of induction beginning at 14 days postpartum. This is also highly similar to the expression profiles observed through clustering in clusters D and E, where the majority of transcripts are induced by 18 and 30 days postpartum, respectively. The spermatocyte- and spermatid-enriched transcript groups were also significantly larger than all the other testicular cell types. This is largely due to the relatively large number of unique genes required to initiate and maintain meiotic activity and produce functional sperm.

These include genes involved in sperm structure, DNA reorganization and condensation, unique metabolism genes, and an assortment of genes involved in sperm-egg binding, motility, and other processes and structures occurring elsewhere in mammalian biology. One of the more intriguing aspects of the analysis of these transcripts was the discovery that many spermatid-enriched or -specific transcripts began being expressed at 14 days postpartum, the period in which spermatocytes, but not necessarily spermatids, are prominent in the testis. This is shown in Figure 3D, which portrays the round spermatid-enriched transcript expression during testis development. Within this group are transcripts encoding zona-pellucida binding proteins, testis-specific serine proteases, and sperm tail structure proteins as well as a large number of EST sequences derived from testis cDNA libraries. Many of these genes would seemingly need to be expressed and subsequently translated toward the end of spermatogenesis but are instead expressed nearly 20 days before the appearance of mature spermatozoa. This observation implies that many of the morphological and biochemical characteristics of spermatozoa are initiated with expression of specific genes during meiosis rather than after the completion of meiotic divisions.

The testis developmental gene expression time course is a step forward in deciphering the regulation of gene expression necessary to elicit testis formation and spermatogenesis. As previously stated, this approach generated data on testis developmental gene expression activity at an exceptional level and, as a result, will allow further inquiries into specific aspects of testis development ultimately controlled by transcriptional events and regulation. Although the general patterns of expression are useful in looking at the testis as an entire tissue, the use of this time course in combination with more specific approaches to studying testis function and other available datasets may in fact be the most vital method of completely describing the testis at the molecular level. Approaches such as knock-out and transgenic models and the establishment and characterization of cell lines to further characterize the functional significance of these changes in gene expression are additional avenues of scientific exploration that can be based on the data within this time course and lead to a better understanding of the functioning of the testis at all levels of inquiry.

ACKNOWLEDGMENTS

The authors would like to thank Derek Pouchnik in the Laboratory for Bioanalysis and Biotechnology I (LBBI) at Washington State University, Pullman, WA, for microarray hybridization and scanning. Thanks also go to Alice Karl and Debra Mitchell for assistance with the mouse colony and Sertoli and myoid cell cultures and Chris Small for critical reading of the manuscript.

REFERENCES


